WHAT IS CLAIMED IS:

1. A method of storing speech information for use in retraining a speech model, the method comprising:

receiving a speech signal;

identifying at least one feature value for each of a set of frames of a speech signal;

decoding the speech signal based on the speech model to identify a sequence of alignment units;

aligning a state of an alignment unit from the sequence of alignment units with a frame in the set of frames of the speech signal; and

before receiving enough frames of the speech signal to begin retraining, adding at least one feature value that is identified for a frame to a feature value sum that is associated with the state that is aligned with the frame.

2. The method of claim 1 wherein the speech signal comprises a single utterance.

3. The method of claim 1 wherein the steps of identifying, decoding, aligning, and adding are repeated for each of a plurality of utterances.

4. The method of claim 3 wherein for each utterance the step of adding to a feature value sum comprises adding to a feature value sum generated from a previous utterance.

5. The method of claim 1 further comprising adding to a frame count associated with a state each time a feature value is added to the feature value sum associated with the state.

6. The method of claim 5 further comprising retraining the speech model based on the feature value sums and the frame counts associated with the states.

7. The method of claim 6 wherein retraining the speech model comprises dividing each state's feature value sum by the state's frame count to form an average value for each state.

8. The method of claim 6 wherein retraining the speech model comprises starting a new computing thread on which the training operations are performed.

9. The method of claim 8 wherein retraining the speech model further comprises updating at least one speech model parameter without locking out the speech model so that the speech model is available for decoding during training.

10. The method of claim 6 further comprising after retraining the speech model repeating the steps of identifying, decoding, aligning, and adding for a new utterance.

11. The method of claim 10 wherein adding to a feature value sum for a state after retraining the

speech model comprises adding to the feature value sum that was used to retrain the model.

12. The method of claim 1 wherein decoding the speech signal further comprises assigning frames to alignment units and wherein aligning comprises aligning the states that form the alignment unit with frames assigned to the alignment unit.

13. The method of claim 12 wherein the alignment unit is a word.

14. The method of claim 5 wherein multiple feature value sums and multiple frame counts are associated with each state.

15. A speech recognition system for recognizing linguistic units in a speech signal, the system comprising:

an acoustic model;

a decoder that uses the acoustic model to identify alignment units in the speech signal;

an aligner that aligns states of the alignment units identified by the decoder with frames of the speech signal;

a dimension sum storage that stores feature dimension sums that are associated with states in the alignment units, at least one state's sums updated before a sufficient number of frames of the speech signal are available to train

the acoustic model, each state's sums updated by summing dimension values from feature vectors assigned to frames aligned with the state; and

a model adapter that uses the feature dimension sums to train the acoustic model after a sufficient number of frames of the speech signal are available.

16. The speech recognition system of claim 15 further comprising a trainer controller that causes the frames of the speech signal to be deleted after the feature dimension sums are formed but before the model adapter trains the acoustic model.

17. The speech recognition system of claim 15 further comprising an initial acoustic model, wherein the model adapter trains the acoustic model by adapting the parameters of the initial acoustic model to form a new version of the acoustic model.

18. The speech recognition system of claim 15 wherein the model adapter is a set of computer-executable instructions that are processed on a different thread from the decoder.

19. The speech recognition system of claim 15 wherein the decoder assigns frames of the speech signal to words and wherein the aligner aligns the frames assigned to a word with the states of the word.

20. A method of aligning frames of a speech signal to states for a sequence of linguistic units, the method comprising:

identifying alignment units corresponding to the sequence of linguistic units and identifying a set of frames that are associated with each alignment unit;

for each alignment unit in the sequence of alignment units:

identifying the states associated with the alignment unit; and

aligning the set of frames associated with the alignment unit by the decoder with the states associated with the alignment unit.

21. The method of claim 20 wherein the method is part of a process of associating feature vectors that represent the speech signal with states of words.

22. The method of claim 21 wherein there is one feature vector per frame and each feature vector comprises a plurality of dimensions.

23. The method of claim 22 wherein the method is used in a process of adapting an acoustic model that further comprises generating a set of dimension sums for each state, each dimension sum being associated with a different dimension of the feature vectors, a dimension sum being formed by summing at least a portion of the values of a respective dimension from all of the feature vectors associated with a state.

24. The method of claim 23 wherein the process of adapting an acoustic model further comprises using the dimension sums to adapt the acoustic model.

25. The method of claim 24 wherein the process of adapting an acoustic model further comprises using the dimension sums to change the parameters of an initial acoustic model to form an adapted acoustic model.

26. A frame alignment system for aligning frames of speech with acoustic states found in alignment units, the alignment system comprising:

a decoder that identifies a sequence of alignment units from a speech signal and associates respective sets of frames of the speech signal with alignment units in the sequence of alignment units;

a trainer controller that identifies acoustic states for the alignment units in the sequence of alignment units; and

an aligner that aligns the acoustic states of an alignment unit with frames in the set of frames associated with the alignment unit.

27. The frame alignment system of claim 26 further comprising an acoustic model that is used by the decoder to identify the sequence of alignment units from the speech signal.

28. The frame alignment system of claim 27 wherein the frame alignment system forms part of a model adaptation system for adapting the acoustic model.

29. The frame alignment system of claim 28 wherein the model adaptation system further comprises a feature extractor that generates a feature vector for each frame of the speech signal, each feature vector comprising a plurality of dimension values for respective dimensions of the feature vector.

30. The frame alignment system of claim 29 wherein the model adaptation system further comprises a dimension sum storage for storing a plurality of dimension sums for each state, each dimension sum being associated with a dimension of the feature vectors and being formed by adding the dimension values for that dimension that are found in the feature vectors that are associated with frames aligned with the state.

31. The frame alignment system of claim 30 wherein the model adaptation system further comprises a model adapter that uses the dimension sums to adapt the acoustic model.